VEHICLE FINGERPRINTING USING DRIVE-BY SOUNDS

V. Cevher[*]*, R. Chellappa*

Center for Automation Research,
University of Maryland,
College Park, MD 20742

A. C. Gurbuz, F. Shah, and J. H. McClellan

Center for Signal and Image Processing,
Georgia Institute of Technology,
Atlanta GA 30332

ABSTRACT

We estimate a vehicle's speed, width, and length by jointly estimating its acoustic wave-pattern using a single passive acoustic sensor that records the vehicle's drive-by noise. The acoustic wave-pattern is estimated using three envelope shape (ES) components, which approximate the shape of the received signal's power envelope. We incorporate the parameters of the ES components along with estimates of the vehicle engine RPM and number of cylinders to form a vehicle profile vector. This vector provides a compressed statistics that can be used for vehicle identification and classification. Vehicle speed estimation and classification results are provided using field data.

1. INTRODUCTION

Estimation of vehicle motion parameters using signals received at passive sensors is a classical signal processing problem [1–6]. When a single passive acoustic sensor is used, wave propagation effects are used to determine the source movements based on the following assumptions that the vehicle $\mathcal{A}$) is a point source [1, 2], $\mathcal{B}$) has stationary signal characteristics that admit a model such as an autoregressive moving average (ARMA) model [2], or $\mathcal{C}$) produces a pure tone [1]. These assumptions are only partially satisfied by vehicles; hence the estimation algorithms based on these assumptions do not perform as expected when they are applied to field data.

When an array of passive acoustic sensors is used, existing approaches in the literature concentrate on the correlation among the multiple microphone signals. Forren and Jaarsma [4] aim to classify vehicles based on their axle detections by exploiting the tire noise generated by vehicles. They use signal correlations among three known microphones under assumption $\mathcal{B}$. However, they do not model any interference effects of the tires as discussed in this paper. Valcarce *et al.* [3] exploits the differential time delays to estimate the speed by assuming $\mathcal{A}$ and $\mathcal{B}$. They use additive Gaussian noise models and obtain biased speed estimates as in [2]. Lo and Ferguson [5] develop a nonlinear least squares method for speed estimation using a quasi-Newton method for computational efficiency. The estimated speed is based on time-delay-of-arrival estimates under assumptions $\mathcal{A}$ and $\mathcal{B}$. Similar to [2, 3], a negative bias in the estimates is also noted in their field tests, which also involve helicopters as targets [5].

In this paper, we provide a power based algorithm for vehicle speed estimation using a single microphone. We describe the spectral and spatial content of vehicle signals and recast the speed estimation problem as a spatial acoustic pattern recognition problem. We calculate the received signal's power envelope and approximate it using three envelope shape (ES) components. The ES components spatially decompose the total vehicle noise into parts that also account for tire interference effects, tire horn effects, and air turbulence effects, which are not considered in the current literature. For estimation, we introduce a vehicle profile vector that characterizes the ES components and also includes classifying vehicle information such as the engine revolutions per minute (RPM) and the number of cylinders. The vehicle profile vector can be thought as a *fingerprint* of the vehicle.

Our motivation for the vehicle profile vector is also the acoustic correspondence problem: given recorded measurements of two vehicles (calibration recordings), we would like to determine, with high confidence, the label of the vehicle when it drives by another control microphone. This problem has applications in distributed sensor networks [7, 8]. The problem becomes complicated when 1) the control microphone has a different distance to the closest point of approach (CPA) of the vehicle, 2) the vehicle is moving with a different speed or moving on a different medium (e.g., gravel as opposed to asphalt), 3) whether or not it is raining (has rained), 4) the vehicle is or was significantly loaded. In this paper, we comment on how we can tackle the correspondence problem using the vehicle profile vector.

2. VEHICLE SIGNAL'S SPECTRAL AND SPATIAL CONTENT

A vehicle's acoustic signal consists of a combination of various noise signals generated by the engine, the tires, the exhaust system, aerodynamic effects, and mechanical effects (e.g., axle rotation, break pads, and suspension). Hence, the spectral content

Prepared through collaborative participation in the Advanced Sensors Consortium sponsored by the U. S. Army Research Laboratory under the Collaborative Technology Alliance Program, Cooperative Agreement DAAD19-01-02-0008.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 NOV 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Vehicle Fingerprinting Using Drive-By Sounds	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Center for Automation Research, University of Maryland, College Park, MD 20742**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002075., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**8**	

of a vehicle's signal includes wideband processes as well as harmonic components. It also has a spatial distribution because the noise sources are at different locations on the vehicle. The mixture weighting of these spectral components at any given location is dependent on the vehicle's speed, whether the vehicle is accelerating, decelerating, turning, and whether the vehicle is in good mechanical condition. In general, one can approximate a vehicle's signal as consisting of four noise components:

Engine Noise: The noise from an internal combustion engine contains a deterministic harmonic train and a stochastic component similar to the human speech [9, 10]. The stochastic component of the engine noise is largely due to the turbulent air flow in the air intake (or intercooler), the engine cooling systems, and the alternator fans. This stochastic component is wideband in nature. The deterministic component is caused by the fuel combustion in the engine cylinders and has more power than the stochastic component. The lowest deterministic tone is called the cylinder fire rate f_0, defined as the firing rate of any one cylinder in the engine. Since each cylinder fires once every two engine revolutions in a four-stroke engine, there is a simple relationship between f_0 and the RPM χ of a vehicle:

$$f_0 = \frac{\chi}{60 \times 2} \text{ Hz}. \tag{1}$$

The strongest tone in the engine noise is called the engine fire rate F_0, and it is related to f_0 in a simple manner:

$$F_0 = f_0 \times p, \tag{2}$$

where p is the number of cylinders in the engine. One can think of F_0 and its integer multiples as the formant frequencies in human speech. The expressions for f_0 and F_0 model the reality quite well; however, small deviations do occur. For example, in modern cars, each cylinder is individually controlled by an engine management system, which might fluctuate f_0 and F_0 to optimize fuel consumption or torque. Hence, in some cases, the locations of f_0- and F_0-harmonics might provide a fingerprint for the specific engine [9].

Car manufacturers try to suppress the engine noise as much as possible for the passengers' comfort inside the vehicle cabin in frequency ranges that the human ears are most sensitive to (1kHz to 4kHz) [10]. In addition, the manufacturers try to suppress the noise levels outside the car as mandated by the federal standards for highway noise (e.g., in the US, see [11,12]). They design quieter engines and also exploit the body of the vehicle to filter the engine noise. To achieve this, the interior of the engine compartment is usually treated with material for acoustical attenuation (the metallic shell also acts as a filter). Hence, in some cases, the engine noise might be stronger on the side and at the very front of the car than other directions, because sound propagation through the axle, the front grill, and the bottom of the engine block cannot be filtered effectively.

Tire Noise The term tire noise is defined as the noise emitted from a rolling tire as a result of its interaction with the road surface. The tire noise is the main source of a vehicle's total noise after 50km/h [13]. It consists of two components: vibrational noise and air noise [14, 15]. The vibrational component is caused by the contact between the tire threads and the pavement texture. Its spectrum is most dominant between $100 - 1000$Hz frequency range. The air noise is generated by the air being sucked-in or forced out of the rubber blocks of a tire and is dominant in the frequency ranges between 1000 and 3000Hz. The actual frequency calculations are complicated by the thread geometry [16].

In the driving direction of the car, the road and the tire forms a geometrical structure that amplifies the noise generated by the tire-road interaction [15, 17, 18]. This effect is called the *horn effect* and has a directional pattern [17]. This amplification results in a strong vehicle tire noise component at far distances in the frequency range $600 - 2000$Hz ([15]: Chapter 7.1.25). The directivity of the horn effect depends on the tire width and radius, the tire shoulders, the tire thread geometry as well as the weight and torque on the tire. Analytical calculations based on these factors are rather difficult, and hence, numerical approaches such as boundary element methods are used to simulate the horn effect for a given tire configuration [17, 18]. Notably, most of the total tire noise power including the horn effect lies between the frequencies of $700 - 1300$Hz with a multi coincidence peak around 1000Hz [15].

Exhaust Noise The exhaust system consists of the exhaust manifold, catalytic converter, resonator, exhaust pipe, muffler, and the tail pipe. The system goes from the engine compartment to the back of the car generating the exhaust noise. Due to the system's spatial distribution, this noise is less prominent in the front of a vehicle. Unlike the engine block noise, the exhaust system noise increases significantly with the engine load. The exhaust noise is also affected by engine turbo/super chargers and after-coolers [19, 20].

Manufacturers use a combination of reactive and absorptive silencers to keep the exhaust noise level down. The exhaust noise has broadband characteristics with most of its power concentrated around the low frequencies. It has the same harmonics frequency structure as the engine and additional tail pipe resonances that occur at fundamental frequency of $f_e = c/(2l)$, where l is the tail pipe length and c is the speed of sound [19, 20].

Air Turbulence Noise Vehicle induced turbulence can become an important factor in the overall perceived loudness of a vehicle as the vehicle speed increases. This noise is due to air flow generated by the boundary layer of the vehicle and is prominent immediately after the vehicle passes by the sensor (by a distinctive *whoosh* sound). The turbulence noise depends on the aerodynamics of the vehicle as well as the ambient wind speed and its orientation [21, 22]. In our problem, we only consider the case when the wind speed is much less than the vehicle speed. For this case, perturbation analysis methods can yield analytical expressions for the mean and the variance of the turbulent velocity components [23]. These expressions may be used to further improve our results in this paper.

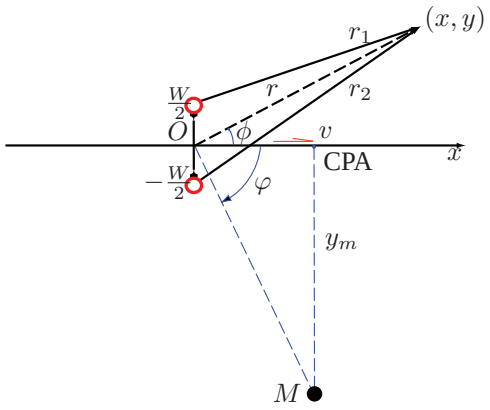


Fig. 1. Dipole geometry. When the dipole sources are correlated, the resulting wave propagation effect on the received signal power is not a superposition of individual monopole effects.

3. INTERFERENCE PHENOMENA

Let $s(t)$ be a zero-mean *i.i.d.* acoustic signal emitted by the monopole source. To simplify the results, we concentrate on the following special case, where the Fourier transform of the source signal is assumed to be bandlimited as follows:

$$|S(\Omega)| = \begin{cases} S, & \Omega_1 \leq \Omega \leq \Omega_2, \mathcal{W} = \Omega_2 - \Omega_1; \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

When the source signal has a spatial extent, it is crucial to consider interference effects while estimating the speed. To demonstrate the interference effects, consider a dipole source moving along the x-axis as illustrated in Fig. 1. In this case, the received signal is the sum of the two source signals that are assumed to be coherent:

$$z[n] = \sum_{i=1,2} \frac{1}{r_i[n]} s\left(\beta_i[n] n F_s - \frac{r_i[n]}{c}\right), \tag{4}$$

where $\beta_i(t)$ $(i = 1, 2)$ is the Doppler shift factor of each monopole source in the dipole. Under the far field assumptions [24], one can approximate $r_i \approx r$ and $\beta_i \approx \beta$ as defined in the monopole source case (Fig. 1).

In the far-field, with the same assumptions for the monopole source, the Fourier transform $Z(\Omega)$ of the signal $z(t)$ can be written as

$$Z(\Omega) \approx \frac{S\left(\frac{\Omega}{\beta[n]}\right)}{\beta[n] r[n]} e^{-j\frac{\Omega}{\beta[n]2c} r_1[n]} \left[1 + e^{-j\frac{\Omega}{\beta[n]2c}(r_2[n] - r_1[n])}\right]. \tag{5}$$

Hence, the received signal bandwidth is modulated as in the monopole case. However, note that the additional term in the brackets in (5) plays a crucial role when we look at the average received signal power:

$$P_z[n] = \frac{S^2 F_s \mathcal{W}}{2\pi\tau} \times \frac{1}{\beta[n] r^2[n]} \times \rho[n], \tag{6}$$

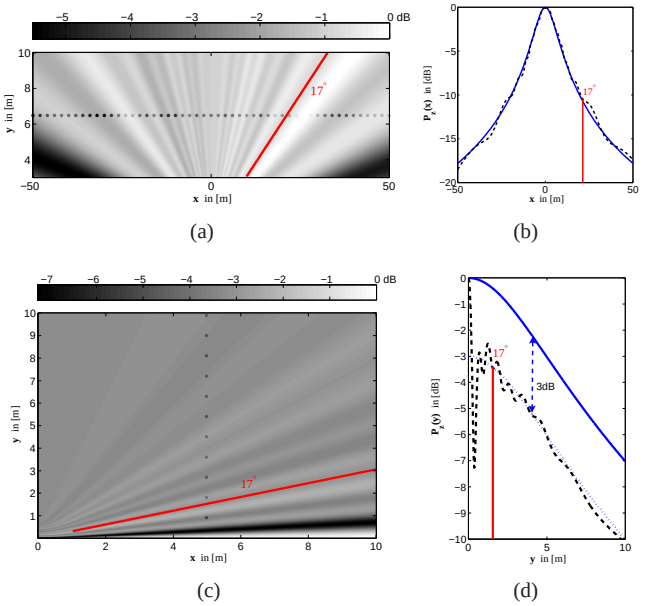


Fig. 2. Interference patterns and power functions along x and y directions for a dipole source with the following parameters: $v = 20$m/s, $W = 1.5$m, $\Omega_1 = 600$Hz, and $\Omega_2 = 2000$Hz. In (b) and (d), the power function is plotted with (dashed line) and without (solid line) the interference term. Note the dramatic effect of the interference term on the power function in (d) along the y direction.

where $\rho[n]$ is called the interference term. When the dipole source signal is baseband, the interference term has the following simpler form:

$$\rho[n] = 1 + \text{sinc}\left(\frac{\mathcal{W}}{\beta[n]c}(r_2[n] - r_1[n])\right). \tag{7}$$

The interference term has a special hyperbolic pattern (Fig. 2). In the far field of the dipole, the interference term is constant along the asymptotes of the hyperbolas defined as $r_2 - r_1 = 2\alpha$ for $\alpha = \frac{W}{2}\sin\phi$. Moreover, it is well-known that the local extremes of the sinc function correspond to its intersections with the cosine function. Hence, a minima and a maxima of the sinc function are on the average half the cosine period away from each other.

4. JOINT ESTIMATION OF SPEED AND SPATIAL ACOUSTIC PATTERNS

4.1. Envelope Shape Components

To determine a vehicle's speed using acoustic observations from a single microphone, we jointly estimate the vehicle's spatial acoustic pattern. In the previous section, we introduced an interference effect that creates a part of the total spatial acoustic pattern. We denote any such component that makes up a vehicle's spatial acoustic pattern as an *envelope shape* (ES) component. Note that,

earlier, we derived the interference effect on the observed acoustic power of the microphone signal *with respect to the source position*. However, in this section, we use the reciprocity theorem and change the reference frame from the moving vehicle to the microphone to derive the ES components [25]. For simplicity, we model the ES components using three piecewise constant functions in dB scale with respect to the microphone bearing φ as illustrated in Fig. 3. We make the connection between the ES components and the received signal power in the next subsection.

The first ES component $\rho_\gamma(\varphi)$ in Fig. 3 models the signal interference due to the front and rear tires, which can be modeled as dipole sources. This component explains the perturbation in the envelope function of the vehicle acoustic drive-by signals around the microphone bearing of $\varphi = 17°$ (Fig. 2). The interference effects before this angle are ignored because they are dominated by the microphone noise. In this ES component, the tire interference decreases between the bearings γ_1 and γ_2, increasing the first component to $\delta_{\gamma,1}$. The angles γ_i are related to the width of the car (dipole separation W) through the interference term.

After the drive-by, the tire interference increases between the bearings γ_2' and γ_1', decreasing the first ES component to $\delta_{\gamma,2}$.

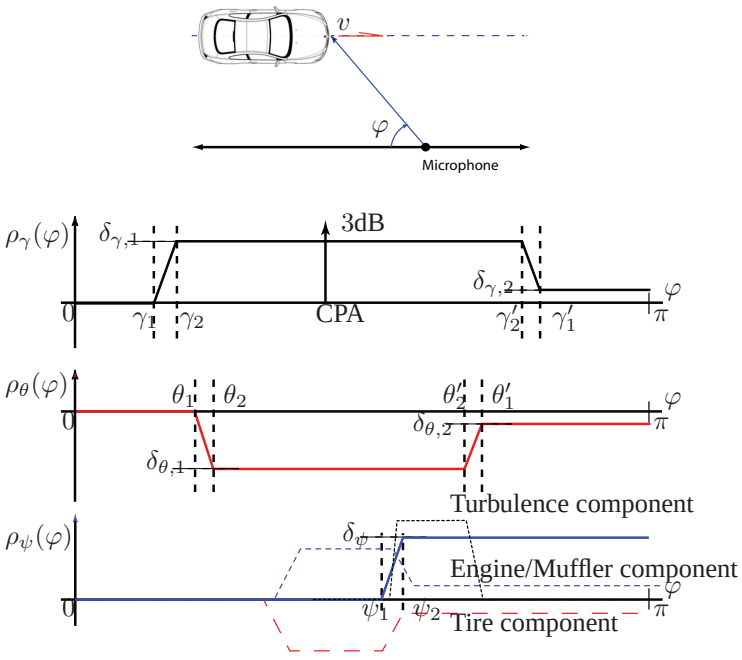


Fig. 3. The microphone bearing reference orientation is defined as the moving direction of the vehicle. Then, a vehicle's spatial acoustic pattern can be approximated by three main components in φ. The first component ρ_γ is due to the signal interference from the front and rear tires. The second component ρ_θ explains the variation as the microphone comes out of the horn-effect area of the tires. The third component ρ_ψ is an approximate component that accounts for a composite engine/exhaust/tire/turbulence noise effect around the vehicle CPA.

The parameter $\delta_{\gamma,2}$ is usually close to zero. We note that the component $\rho_\gamma(\varphi)$ varies in a nonsymmetric fashion with respect to φ. The asymmetry is due to the movement of the car: because of the reference frame change, any angle defined in the vehicle reference frame, denoted by ϕ, is related to the angles in the microphone frame, denoted φ, through an aberration relation [26]:

$$\tan\frac{\varphi}{2} = \sqrt{\frac{1+v/c}{1-v/c}}\tan\frac{\phi}{2}, \tag{8}$$

where the sign of the speed terms flip after the CPA. Hence, by assuming a symmetric interference pattern for the front and rear tires of the car based on constant car width, one can relate the following angle parameters:

$$\tan\frac{\kappa'}{2} = \pi - \left(\frac{1-v/c}{1+v/c}\right)\tan\frac{\kappa}{2}, \tag{9}$$

where $\kappa = \gamma_1$ and γ_2.

The second ES component $\rho_\theta(\varphi)$ is due to the horn effect, which was explained in Sect. 2. In the observed signal envelope, at the microphone bearing θ_1 the horn effect amplification of the farthest front tire from the microphone starts to go down until the bearing θ_2 to $\delta_{\theta,2}$, when the horn effect of the closest rear tire to the microphone also drops. The differential angle $\theta_2 - \theta_1$ is a very good indicator of the vehicle length, which can be used to compare the relative sizes of vehicles. To convert the angle difference into actual size, we use the following approximate relationship

$$\frac{y_m}{\cos\theta_1} - L = \frac{y_m}{\cos\theta_2}, \tag{10}$$

where L is the car length. Geometrically, θ_2 is the microphone bearing of the front of the car after the car fully passes the line defined by the bearing θ_1.

After the CPA, the horn effect amplifies the tire noise between bearings θ_2' and θ_1', which are related to θ_1 and θ_2 also by the abberration relation (9) (i.e., $\kappa = \theta_i$). The final level of the tire noise component $\delta_{\theta,2}$ is usually different than 0dB, because the rear tire curvature is different from the front tire curvature due to the torque on the tire. Any imbalance of the weight ratio on the front and rear tires also causes $\delta_{\theta,2}$ to deviate from 0dB.

Finally, *the third ES component* $\rho_\psi(\varphi)$ is a composite component that incorporates (i) engine noise, (ii) exhaust system noise, (iii) interference pattern of the tires on the side of the car, and (iv) the noise caused by the air turbulence. To keep the number of ES components manageable, we approximate the composite interference pattern as a step function that rises from 0dB to δ_ψ between bearings ψ_1 and ψ_2. When this approximation becomes poor, $\delta_{\theta,2}$ of the second ES component $\rho_\theta(\varphi)$ compensates. We found that the angle difference $\psi_2 - \psi_1$ is also an indicator of the vehicle length. Hence, (10) is also used to relate the angles in the third interference component to the car length L.

4.2. Vehicle Profile Vector

To jointly determine the speed and the vehicle's spatial acoustic pattern, we use the vehicle profile vector $\boldsymbol{\lambda}$, which is defined as follows:

$$\boldsymbol{\lambda} = \left[\begin{array}{cccc} \boldsymbol{\lambda}_v & \boldsymbol{\lambda}_\varphi & \boldsymbol{\lambda}_\delta & \boldsymbol{\lambda}_f \end{array} \right], \text{ where} \tag{11}$$

$$\begin{aligned} \boldsymbol{\lambda}_v &= \left[\begin{array}{cccc} S & v & W & L \end{array} \right], \boldsymbol{\lambda}_\varphi = \left[\begin{array}{cccc} \varphi_0 & \gamma_1 & \theta_1 & \psi_1 \end{array} \right], \\ \boldsymbol{\lambda}_\delta &= \left[\begin{array}{ccccc} \delta_{\gamma,1} & \delta_{\gamma,2} & \delta_{\theta,1} & \delta_{\theta,2} & \delta_\psi \end{array} \right], \text{ and } \boldsymbol{\lambda}_f = \left[\begin{array}{cc} \chi & p \end{array} \right]. \end{aligned} \tag{12}$$

The vector $\boldsymbol{\lambda}_v$ consists of the physical parameters of the vehicle such as the loudness S, speed v, car width W, and the car length L. The vector $\boldsymbol{\lambda}_\varphi$ has the initial vehicle bearing φ_0 and the angles that define the ES components along with $\boldsymbol{\lambda}_\delta$, which contains the amplitude attenuations and amplifications for the ES components. Lastly, the vector $\boldsymbol{\lambda}_f$ has the RPM χ and the number of cylinders p of the vehicle. The profile vector $\boldsymbol{\lambda}$ can be viewed as a *fingerprint* of the vehicle and can be used for appearance-based tracking and classification.

4.3. Amplitude Observations

In this section, we derive a relationship between the vehicle profile vector $\boldsymbol{\lambda}$ and the square-root of the average signal power, which we will denote as the power envelope. This relationship is used to determine the vehicle profile vector using standard maximum likelihood estimation techniques.

We define the power envelope function by using τ-discrete samples of $z[n]$ as follows:

$$\mathcal{E}[n_\tau] = \mathcal{E}(t)\Big|_{t=\frac{n_\tau \tau}{F_s}} = \sqrt{P_z[n\tau]}, \tag{13}$$

where subscript τ under the sample index n implies that the samples of the continuous function are calculated at every τ/F_s second. The parameter τ is chosen so that the DFT coefficients used to calculate the power function at τ-samples apart are statistically uncorrelated, and hence, each sample of $\mathcal{E}[n_\tau]$ ($n_\tau = 0, 1, \ldots, N_\tau - 1$) is also statistically uncorrelated of the others.

Assuming that the noise acting on the microphone signal $z[n]$ is zero-mean additive white Gaussian noise with variance σ_u^2, we relate the envelope observations to the vehicle profile vector as follows:

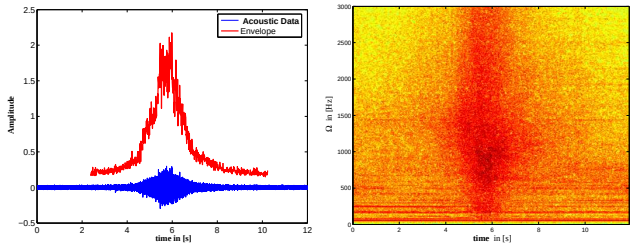
$$\begin{aligned} \mathcal{E}^2[n_\tau] &\approx \mathcal{A}_{\boldsymbol{\lambda}}[n_\tau] e^{2m_\tau} + \frac{\sigma_u^2}{\tau} w_\tau, \\ \mathcal{A}_{\boldsymbol{\lambda}}[n_\tau] &= \frac{C}{\beta[n_\tau] r^2[n_\tau]} \prod_{i=\gamma,\theta,\psi} 10^{\rho_i(\varphi[n_\tau])/10}, \quad C = \frac{S^2 F_s \mathcal{W}}{2\pi\tau} \end{aligned} \tag{14}$$

where $\mathcal{A}_{\boldsymbol{\lambda}}[n_\tau]$ is the directional power variation, e^{m_τ} is an *i.i.d.* multiplicative noise on the signal amplitude ($m \sim \mathcal{N}\left(0, \sigma_m^2\right)$), w_τ is an *i.i.d.* additive Chi2$_\tau$ noise (chi-squared distribution with τ degrees of freedom) that is also independent of m_τ, and

$$\begin{aligned} \beta[n_\tau] &= 1 - \frac{v}{c} \cos\varphi[n_\tau], \\ r[n_\tau] &= \sqrt{(v\tau/F_s)^2 + r^2[n_\tau - 1] - 2\,(v\tau/F_s)\, r[n_\tau - 1] \cos\varphi[n_\tau - 1]}, \\ \varphi[n_\tau] &= \varphi[n_\tau - 1] + \sin^{-1}\left(\frac{v\tau}{F_s r[n_\tau]} \sin\varphi[n_\tau - 1] \right), \\ r[0] &= y_m \sec\varphi_0. \end{aligned} \tag{15}$$

In (15), we implicitly assume that the constant velocity motion equations are not violated even though the reference frame is changed from the moving vehicle to the stationary microphone.

Let $\boldsymbol{\mathcal{E}} = \left[\begin{array}{ccccc} \mathcal{E}[0] & \ldots & \mathcal{E}[n_\tau] & \ldots & \mathcal{E}[N_\tau - 1] \end{array} \right]$ denote the aggregate envelope observations. Then, the observation likelihood, given the vehicle profile vector as well as the noise variances σ_m^2 and σ_u^2, can be determined by a straightforward bivariate transformation followed by marginalization [27]. Unfortunately, the marginalized data likelihood does not have a closed form solution and needs to be evaluated numerically. Moreover, note that the noise variances σ_m^2 and σ_u^2 have to be determined for the evaluation of the likelihood. A joint estimation of the vehicle profile vector and the noise variances can be done. However, this increases the numerical complexity. In theory, these noise variances can be treated as nuisance parameters and can be integrated out using reference priors [28]. In practice, they further increase the necessity of numerical integration.



(a) (b)

Fig. 4. Drive-by test by a 6-cylinder Chevy Impala moving with 18.7m/s at an approximate distance of 5.8m. (a). The acoustic signal was sampled at $F_s = 48$kHz. The power envelope $\boldsymbol{\mathcal{E}}$ is calculated with $\tau = 480$. In the figure, $\sqrt{\tau}\boldsymbol{\mathcal{E}}$ is plotted to emphasize the variation. There is an asymmetry in the envelope estimates that can be explained by the ES components. (b). The spectral content of the acoustic signal. Note the strong interference at 60Hz. The tire-noise spectrum, which is concentrated around $700 - 1300$Hz, does not exhibit a frequency modulation pattern as predicted by the theory.

4.4. Frequency Observations

The spectral content of a vehicle exhibits directional variation, making it difficult to use the frequency modulation effects of the

vehicle motion to determine speed. We emphasize that this directional variation is not due to the motion of the vehicle but it is due to tire noise effects, which are stochastic in nature as discussed in Sect. 2. The useable frequency tracks for speed estimation are generated by the engine because the frequency modulation effects can be observed in the deterministic component of the engine noise. These deterministic engine frequencies span the $0-250$Hz range at nominal RPM's (e.g., Fig. 4(b)). At moderate vehicle speeds ($30-50$mph), the full Doppler shift swings F_0 approximately %6, also corresponding to an RPM change of the same amount ($\Delta\chi \approx 200$). Hence, if a driver changes the car's RPM by 50 during the vehicle drive-by, there will be a %25 error in F_0 when one assumes a constant frequency source. We emphasize that this RPM change is unnoticeable on the dashboard of the vehicle and is likely to happen. On the other hand, the effect of the same RPM change on the total car loudness is negligible.

Therefore, determining a probability density function for the vehicle profile vector by fitting a Doppler shift function to the engine and tire frequency tracks is an unreliable approach. For example, in [2], the speed estimation was performed using an autoregressive modeling of the acoustic signals under a point source assumption. It was concluded that the Doppler-based speed estimation on the source frequencies does not perform well with field-data [2]. It was also concluded that, with the same source assumptions, the envelope measurements yield improved speed estimates than the frequency measurements; however the speed estimates are nonetheless biased. A possible reason of this bias mentioned in [2] and also observed in [3–5] is the additive Gaussian model as opposed to the multiplicative noise model that we employ in this paper.

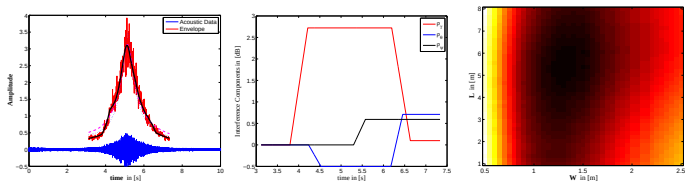
On the other hand, the spectral harmonic content can be used to determine $\boldsymbol{\lambda}_f$ of the vehicle profile vector. Moreover, conditioned on $\boldsymbol{\lambda}_v$ estimate, it possible to further refine $\boldsymbol{\lambda}_f$ by compensating for the Doppler shifts. The number of cylinders p is usually the most elusive to estimate because the body of a vehicle may also act as a filter to directionally suppress the frequency at the engine fire rate F_0. Hence, it is rather easy to incorrectly estimate the number of cylinders of a vehicle because the strongest frequency is not necessarily F_0. If a characterization can be done, which is applicable to the vehicles of interest, the number of cylinders can also be estimated robustly. Estimation of χ can be done accurately using harmonic analysis methods [29]. In our estimation, we use the power spectral density of the acoustic signal to determine $\boldsymbol{\lambda}_f$. Details can be found in [29].

5. EXPERIMENTS

To demonstrate the ideas, data was collected with $F_s = 48$KHz at a two-way street with an omnidirectional microphone, emplaced 1.5m off the ground on a pole at the sidewalk. The distance of the bottom of the microphone pole to the center of the street is 7.4m. A video camera is used to establish the ground truth and identify the vehicles in the test.

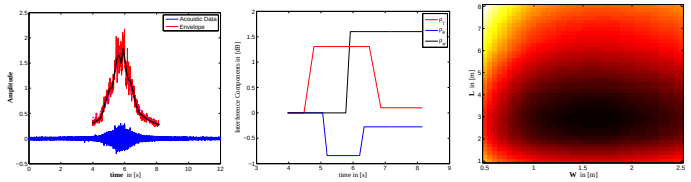
5.1. Vehicle Profiling

Table 1 lists the results of the vehicle speed estimates obtained by three different methods using $\tau = 480$ samples: 1) the full vehicle profile vector using (14), 2) only $\boldsymbol{\lambda}_v$ using (14), and 3) only $\boldsymbol{\lambda}_v$ using the constant velocity motion model on a point source as in [1, 2]. It is seen that the estimates of [2] are improved by incorporating the multiplicative noise model on the signal envelope. Estimation using the ES components yields the best estimates (Figs. 5 and 6). We also estimated the number of cylinders and the RPM of the vehicles by using the methods in [29]. The number of cylinder estimates p are estimated by compensating for the microphone spectral characteristics. However, there was no compensation for any vehicle directional variation.



(a) (b) (c)

Fig. 5. **Ford F150**. (a) Estimated envelope by the interference components is shown with the solid line. The dashed line and the dotted line belong to the additive and multiplicative noise model results, respectively. (b) Estimated ES components are shown. According to the ES components, the vehicle is louder in the rear than front. (c) Estimated joint distribution of the vehicle dimensions around the solution.



(a) (b) (c)

Fig. 6. **Chevy Impala**. (a) Observed envelope exhibits significant variations. The interference components (solid line) adequately explains the variations. (b) Estimated ES components are shown. Parameter δ_ψ is relatively larger than the other components indicating significant air noise. (c) Estimated joint distribution of the vehicle dimensions around the solution.

5.2. Vehicle Classification Results

The vehicle profile vector provides a natural basis for classifying vehicles. Figures 7(a) and (b) show that the vehicles can be separated into two classes based on their length and size. Note that the estimated vehicle lengths are not exact vehicle lengths; however, they can separate compact cars from SUV's or trucks. Figure 7(c)

Table 1. Field Test Results

Ground Truth			Estimation using $\boldsymbol{\lambda}$						Estimation using $\boldsymbol{\lambda}_v^{\S}$		Estimation using $\boldsymbol{\lambda}_v^{\dagger}$	
Vehicle	y_m	$v$camera	v	C	L	$W^{\ddagger}$	χ	$p^{\natural}$	v	C	v	C
Ford F150	6.3	17.54m/s	17.86m/s	12.60	5.38m	1.30m	3038	8	28.00m/s	24.08	21.39m/s	21.27
Chevy Impala	5.8	18.68m/s	18.60m/s	9.23	2.58m	1.75m	3300	6	18.29m/s	11.55	15.05m/s	10.90
Honda Accord	4.3	16.74m/s	14.44m/s	6.86	3.28m	1.40m	3074	6	17.34m/s	10.17	14.49m/s	9.67
Nissan Maxima*	4.6	13.32m/s	13.20m/s	12.45	3.28m	1.50m	3825	$6'$	14.23m/s	14.86	14.27m/s	14.49
Nissan Maxima*	4.1	4.14m/s	4.49m/s	6.34	2.58m	1.50m	3150	4	4.75m/s	9.90	3.46m/s	9.20
Isuzu Rodeo	8.1	13.44m/s	13.89m/s	7.87	5.20m	1.35m	3450	6	11.32m/s	7.97	11.79m/s	7.95
Mercedes E	8.1	13.94m/s	13.80m/s	7.68	2.93m	1.50m	3075	6	15.51m/s	10.47	11.78m/s	9.93
Volvo 850 SW	8.1	14.11m/s	14.69m/s	9.60	3.10m	1.40m	2250	$10'$	12.93m/s	9.01	11.22m/s	8.63
Nissan Frontier	4.3	17.56m/s	17.84m/s	9.31	4.85m	1.40m	2625	6	17.02m/s	9.92	17.56m/s	9.74
VW Passat	5.1	11.66m/s	11.58m/s	6.06	2.75m	1.80m	1950	6	8.58m/s	6.06	8.66m/s	6.11
Error STD			0.8246m/s						3.7203m/s		2.2627m/s	
Error STD¶			0.2777m/s						1.5154m/s		1.5126m/s	
Bias			-0.0735m/s						0.6845m/s		-1.1458m/s	
Bias¶			0.1737m/s						-0.4017m/s		-1.7013m/s	

§ Using the method in [2] with multiplicative noise model introduced here.

† Using the method in [2] without any change.

‡ A fixed bandwidth of $\mathcal{W} = 600$Hz is used to determine the car widths.

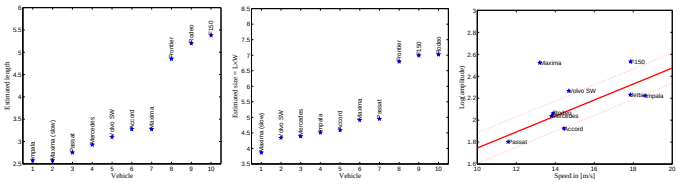
♮ Estimated by finding the frequency F_0 with the maximum power spectral density between frequencies 85-210Hz and then dividing F_0 by the CFR f_0 estimate [29].

′ Incorrectly estimated. The actual values are 4 (Maxima) and 5 (Volvo).

* Same vehicle.

¶ Calculated by removing one outlier in each method.

also illustrates that it is possible to identify loud vehicles such as vehicles with mechanical problems or heavily loaded SUV's or pick-up trucks, which are expected to be louder than usual. This classification is based on the fact that the loudness of the vehicle has a certain functional distribution as indicated in [14, 15]. Hence, given two similar vehicles, it may be possible to identify if one of them is heavily loaded or has mechanical problems even if they move at different speeds.



(a) (b) (c)

Fig. 7. (a) Estimated vehicle lengths are compared. There is a clear separation between compact cars and large vehicles. (b) Estimated vehicle sizes are compared. (c) Logarithm of the vehicle signal amplitudes are plotted with respect to their speed. There is a linear trend in the plot as also indicated by [14, 15]. The solid line represents a least squares fit to the data without Nissan Maxima. The dotted lines are one standard deviation away from the mean. Nissan Maxima is louder than the other cars because the vehicle has mechanical problems.

6. CONCLUSIONS

We presented a method to determine a vehicle's speed via its acoustic drive-by sounds recorded at a microphone, by formulating the problem as a joint speed and acoustic pattern estimation problem. We achieve this estimation using a vector that profiles the directional variation of the vehicle acoustic pattern. The vehicle profiles vector enables a signal processor to better address the vehicle correspondence problem since the vehicle profile vector provides unbiased speed and loudness estimates as well as vehicle dimensions. It also generates better discriminative features that are compressed into a 15-dimensional space. Parameters $\boldsymbol{\lambda}_v$ and $\boldsymbol{\lambda}_f$ of the vehicle profile vector can improve the confidence of the correspondence matches, whereas their compression decreases communication among a calibration microphone and a control microphone. However, as usual, given the difficulty of the correspondence problem, one should not expect superlative performance for all cases even with the vehicle profile vector.

While determining the vehicle speed, we relied on the signal power calculations and argued that the signal frequency information (Doppler) was not useful. On the other hand, when an array of microphones is available, one can also infer from the phase of the received acoustic data across the array. In this case, we expect that the performance should improve more than that one would expect to obtain from multiple independent amplitude observations. We envision that when multiple vehicles are present, the array can provide the acoustic steering necessary to remove the cocktail party effect on the ES components. Hence, the approaches in the literature can be improved to obtain unbiased speed estimates when an array is used.

7. REFERENCES

[1] B. G. Quinn, "Doppler speed and range estimation using frequency and amplitude estimates," *J. Acoust. Soc. Am.*, vol. 98, pp. 2560–2566, November 1995.

[2] C. Couvreur and Y. Bresler, "Doppler-based motion estimation for wide-band sources from single passive sensor measurements," in *ICASSP 1997*, 21–24 April 1997.

[3] R. Lopez-Valcarce, C. Mosquera, and R. Perez-Gonzalez, "Estimation of road vehicle speed using two omnidirectional microphones: a maximum likleihood approach,"

EURASIP Journal on Applied Signal Processing, pp. 1059–1077, 2004.

[4] J. F. Forren and D. Jaarsma, "Traffic monitoring by tire noise," in *Proc. IEEE Conf. on Intelligent Transportation System*, Boston, MA, Nov 1997, pp. 177–182.

[5] K. W. Lo and B. G. Ferguson, "Broadband passive acoustic technique for target motion parameter estimation," *IEEE Trans. on Aero. and Elec. Syst.*, vol. 36, pp. 163–175, 2000.

[6] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall, 1993.

[7] S. Kumar, F. Zhao, and D. Shepherd, "Collaborative signal and information processing in microsensor networks," *IEEE Signal Processing Magazine*, vol. 19, pp. 13–14, 2002.

[8] M. Duarte and Y-H. Hu, "Vehicle classification in distributed sensor networks," *Journal of Parallel and Distributed Computing*, vol. 64, pp. 826–838, 2004.

[9] S. A. Amman and M. Das, "An efficient technique for modeling and synthesis of automotive engine sounds," *IEEE Transactions on Industrial Electronics*, vol. 48, pp. 225–234, February 2001.

[10] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*, Prentice-Hall, 1978.

[11] Senate Public Works Committee, *Noise Pollution and Abatement Act of 1972*, S. Rep. No. 1160, 92nd Cong. Second session, 1972.

[12] 86 Stat. 1234 Public Law No. 92-574, *Noise Pollution and Abatement Act of 1972*, codification amended at 42 U.S.C. 4901-4918, 1988.

[13] U. Sandberg, "Tyre/road noise - myths and realities," in *Proceedings of the 2001 International Congress and Exhibition on Noise Control Engineering*, Hague, Netherlands, 27–30 August 2001.

[14] Road Directorate-Ministry of Transport, "Noise reducing pavements," Tech. Rep. 141, Road Directorate, Danish Road Institute, April 2005.

[15] U. Sandberg and A. J. Ejsmont, *Tyre/road noise reference book*, Infomex, SE-59040 Kisa, Sweden, 2002.

[16] P. Andersson, *High frequency tyre vibrations*, Ph.D. thesis, Dept. of Applied Acoustics, Chalmers Univ. of Tech., Gothenburg, Sweden, 2002.

[17] R. A. G. Graf, C. Y. Kuo, A. P. Dowling, and W. R. Graham, "On the horn effect of a tyre/road interfacepart I: experiment and computation," *Journal of Sound and Vibration*, vol. 256, pp. 417–431, 2002.

[18] C. Y. Kuo, R. A. G. Graf, A. P. Dowling, and W. R. Graham, "On the horn effect of a tyre/road interfacepart II: asymptotic theories," *Journal of Sound and Vibration*, vol. 256, pp. 433–445, 2002.

[19] S. M. Kuo and D. R. Morgan, "Active noise control: a tutorial review," *Proceedings of the IEEE*, vol. 87, pp. 973–973, 1999.

[20] J. G. Lilly, "Engine exhaust noise control," available online at *http://www.ashraeregion7.org*.

[21] R. E. Eskridge and J. C. R. Hunt, "Highway modeling. Part I: prediction of velocity and turbulence fields in the wake of vehicles," *American Meteorological Society*, vol. 79, pp. 387–400, 1979.

[22] N. Sarigul-Klijn, D. Dietz, D. Karnopp, and J. Dummer, "A computational aeroacoustic method for near and far field vehicle noise predictions," *The American Institute of Aeronautics and Astronautics Inc.*, 2001.

[23] P. Kastner-Klein, R. Berkowicz, and E. J. Plate, "Modelling of vehicle induced turbulence in air pollution studies for streets," in *5th Workshop on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes*, Rhodes, Greece, 18–21 May 1998.

[24] L. J. Ziomek, "Three necessary conditions for the validity of the Fresnel pahse approximation for the near-field beam pattern of an aperture," *IEEE Journal of Oceanic Engineering*, vol. 18, pp. 73–76, 1993.

[25] Lord Rayleigh, *The Theory of Sound*, Macmillan, London, 1877.

[26] E. Einstein, "On the electrodynamics of moving bodies," *Annalen der Physik*, vol. 17:891, 1905.

[27] A. Papoulis and S. U. Pillai, *Probability, random variables and stochastic processes*, McGraw Hill, 2002.

[28] J. O. Berger and J. M. Bernardo, "On the development of reference priors," *Bayesian Statistics*, vol. 4, pp. 35–60, 1992.

[29] B. Boashash and P. O'Shea, "A methodology for detection and classification of some underwater acoustic signals using time-frequency analysis," *IEEE Trans. on ASSP*, vol. 38, pp. 1829–1841, Nov 1990.